PAPER Special Section on Signal Processing

Differential Energy Based Watermarking Algorithm Using Wavelet Tree Group Modulation (WTGM) and Human Visual System

Min-Jen TSAI[†a)], *Member and* **Chang-Hsing SHEN**[†], *Nonmember*

SUMMARY Wavelet tree based watermarking algorithms are using the wavelet coefficient energy difference for copyright protection and ownership verification. WTQ (Wavelet Tree Quantization) algorithm is the representative technique using energy difference for watermarking. According to the cryptanalysis on WTQ, the watermark embedded in the protected image can be removed successfully. In this paper, we present a novel differential energy watermarking algorithm based on the wavelet tree group modulation structure, i.e. WTGM (Wavelet Tree Group Modulation). The wavelet coefficients of host image are divided into disjoint super trees (each super tree containing two sub-super trees). The watermark is embedded in the relatively high-frequency components using the group strategy such that energies of sub-super trees are close. The employment of wavelet tree structure, sum-of-subsets and positive/negative modulation effectively improve the drawbacks of the WTQ scheme for its insecurity. The integration of the HVS (Human Visual System) for WTGM provides a better visual effect of the watermarked image. The experimental results demonstrate the effectiveness of our algorithm in terms of robustness and imperceptibility.

key words: *copyright protection, Human Visual System, image watermarking, wavelet, wavelet tree quantization*

1. Introduction

Digital media files can be easily copied and distributed without any reduction in quality. As a result, digital media files are being widely distributed on the Internet today, through both authorized and unauthorized distribution channels. Piracy is a concern when security measures are not in place to protect content.

Conventional cryptographic systems permit only valid principals (key holders) access to encrypted data. Once such digital data are decrypted, there is no way to track their reproductions or retransmissions. Over the last decade, digital watermarking has been presented to complement cryptographic protection mechanisms. Invisible watermarks can be broadly classified into three types, i.e. robust, fragile (or semi-fragile) and captioning watermarks [1], [2]. Robust watermarks are generally used for copyright protection and ownership verification as they are robust to nearly all kinds of image processing attacks. Fragile or semi-fragile watermarks are mainly applied to content authentication and integrity attestation as they are fragile to most modifications. Captioning watermarks are usually used for side information conveyance, which are required to convey more information than robust watermarks do.

Manuscript received November 12, 2007.
Manuscript revised March 10, 2008.

[†]The authors are with Institute of Information Management, National Chiao Tung University, Hsing-Chu, 300 Taiwan.
a) E-mail: mjtsai@cc.nctu.edu.tw

DOI: 10.1093/ietfec/e91–a.8.1961

Cox et al. [1] proposed a global DCT-based spread spectrum approach to hide watermarks. The frequency domain of the image or sound is viewed as a communication channel, and correspondingly, the watermark is viewed as a signal that is transmitted through it. The watermark is spread over very many frequency bins so that the energy in any one bin is very small and certainly undetectable. Langelaar and Lagendijk [3] introduced the DEW (Differential Energy Watermarking) algorithm for JPEG/MPEG streams in the DCT domain. The DEW algorithm embeds label bits (the watermark) by selectively discarding high frequency DCT coefficients in certain image regions. Wang and Lin [4] introduced the philosophy of WTQ (Wavelet Tree Quantization) in the DWT domain. The wavelet coefficients are grouped into so-called super trees. The wavelet-tree-based watermarking algorithm embeds watermark bits by selectively quantizing super trees.

Whether the security of watermarking algorithms can be preserved if the details about algorithms are released is always a controversial issue among watermarking researchers. However, the algorithm will be known to the attacker as it is accepted in the field of cryptology [5]. Das, Maitra and Mitra had presented a successful cryptanalysis against the DEW scheme in [5]. There is a need to analyze each of the popular watermarking algorithms individually and to check whether customized attacks can be mounted to highlight the weakness of the individual watermarking algorithm itself.

In this paper, we first introduce the WTQ scheme and then explain how this watermarking algorithm can be attacked by cryptanalysis. Based on the motivation to improve the security robustness of WTQ, we present a differential energy watermarking algorithm based on the wavelet tree group modulation structure, i.e. WTGM (Wavelet Tree Group Modulation). The usage of group modulation makes the proposed watermarking algorithm robust against common signal processing attacks and results in a better detector response. With the characteristic of the wavelet tree structure throughout large spatial regions, it is more robust against geometric distortions. The employment of sum-of-subsets makes the proposed watermarking algorithm more robust against general cryptanalysis. In addition, the consideration to the CSF (Contrast Sensitive Function) and NVF (Noise Visibility Function) of the HVS (Human Visual System) provides a better visual effect of the watermarked image.

The remainder of this paper is organized as follows. In

Sect. 2, the WTQ scheme is briefly explained and its vulnerability is reviewed in Sect. 3. In Sect. 4, the proposed WTGM watermarking method is described in details. The experimental results and conclusions are given in Sects. 5 and 6, respectively.

2. WTQ Scheme

The WTQ scheme is a wavelet tree based blind watermarking scheme. Interested reader can refer Ref. [4] for detailed information and we briefly explain the operations:

2.1 Group the Super Tree

In the WTQ scheme, a pair of super trees is used to record one watermark bit, and the watermark is a binary PN sequence of {1, −1}. Before recording all watermark bits, we should perform 4-level DWT as shown in Fig. 1(a), and collocate coefficients in $C_{i,j}$, where $i = \{2, 3, 4\}$ and $j = \{1, 2, 3\}$, to form the groups in Fig. 1(b). A group has 21 coefficients: 1 coefficient from level-4, 4 coefficients from level-3, and 16 coefficients from level-2. After grouping, two groups are randomly combined to become a super tree and there are 42 coefficients in every super tree.

While all super trees are grouped, WTQ starts to embed the watermark bits. Here we pair two super trees with a secret seed for quantization operation, so one super tree of a pair will be quantized for recording the information of the corresponding watermark bit. For example, if the current watermark bit is 1, we will quantize the left super tree in the corresponding pair. Otherwise, the right one will be quantized.

In order to perform quantization, all super trees will be transformed to "bit-plane" form first.

2.2 Form the Bit-Plane for Quantization

A super tree will be transformed to the bit-plane format as shown in Fig. 1(c) for calculating the scope of bits which will be removed according to the reference error. In Fig. 1(c), the energy of the gray area is the same with or bigger than the reference error. When we have found the scope, bits in the gray area will be discarded.

In WTQ, every watermark bit is recorded with a tree-pair. If we want to extract the watermark, both trees in a tree-pair will be checked and then we will find out which tree is quantized. Thus, according to the quantization pattern, we can get the information of the current watermark bit.

3. Cryptanalysis on WTQ

In [6], Das and Maitra claimed that every existing watermarking algorithm should be tested as a cryptographic model by cryptanalysis. Therefore, Das and Maitra put the cryptanalytic techniques on the WTQ scheme, and found out the weakness. As the paper says, "Knowledge of groups, which is image dependent, but not dependent on secret seed,

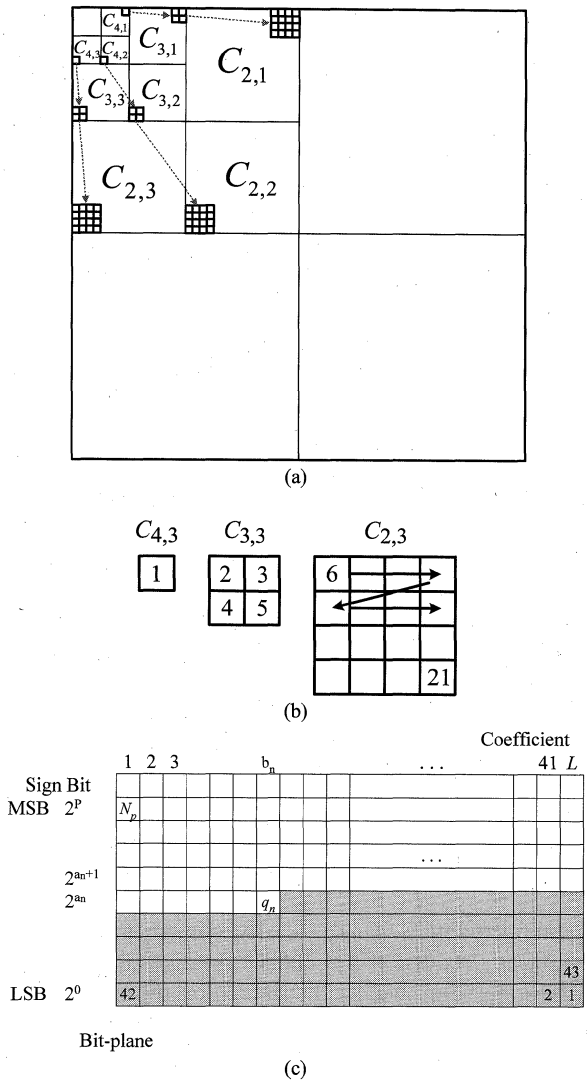


Fig. 1 Illustrations of the wavelet transform and groups of wavelet coefficients. (a) A four-level wavelet transform and its subbands. The coefficients correspond to the same spatial location are grouped together. (b) The 21 coefficients of a wavelet tree. (c) Binary representation of the coefficients in the nth super tree. The shaded area denotes the bits to be discarded.

is enough for successful removal of correlation." Although the information of constructing a super tree is unknown, but we know that every super tree is obtained with two groups. In order to destroy the watermarking information in super tree, we can use indirect technique through the knowledge of groups.

The cryptanalysis on WTQ can be divided into tree steps: identification of quantized and non-quantized groups, estimate of reference error, and quantization of non-quantized groups.

3.1 Identification of Quantized and Non-quantized Groups

All groups will be transformed into bit-plane format for identification. The principle is calculating the energy of last

rows. In our simulation, last two rows are used. If the bits of last rows in current group are almost empty, we can assume this group as a quantized group. Otherwise, this group is a non-quantized group.

3.2 Estimate of Reference Error

After identification, we take the set of quantized groups for estimate. First, we calculate the quantization error of all groups in the set, and find out that the energy removed in every group is almost ε'. Thus, ε' is the estimated reference error.

3.3 Quantization of Non-quantized Groups

When reference error has been estimated, the set of non-quantized groups will be quantized using this estimated reference error. After this step, all groups are almost quantized.

In WTQ, every watermark bit is recorded by quantizing only one tree in a pair. Making all groups quantized means making all super trees quantized because a super tree is merged with two groups. Thus, if all trees are quantized, the difference caused by quantization between two trees in a pair will be eliminated. As the difference between both trees declines, it is difficult for the detector to extract the watermark bit accurately.

According to our simulation of the cryptanalysis attack for WTQ, the unquantized bitplane could be successfully identified and the last two rows could be removed. Therefore, the watermark will be removed even without the reference error estimation. Therefore, WTQ is not secure enough for digital watermarking in principle.

4. Designs of WTGM Algorithm

There are several issues need to be addressed if the energy difference will be applied for the wavelet based watermarking scheme. The first is the choice of the tree structure. How many levels should the image to be decomposed and achieve the robustness and the scalability of the watermark? The second is how to balance the robustness and the fidelity of the image on a designed energy differential watermarking algorithm? The third is how to maximize the detector response in order to render a better performance. The fourth is the security of the watermarking algorithm, the most important issue to be addressed. To resolve those mentioned issues, the same pyramidal decomposition is applied in WTGM. In addition, the idea of sum-of-subsets [5] for selecting supertrees is adopted to securely embed the watermark. Instead of the bit plane quantization for watermark embedding, the usage of positive/negative modulation will effectively render a better detector response. Also, the consideration to the CSF and NVF of the HVS provides a better visual effect and imperceptibility. The details will be explained next.

4.1 PM (Positive Modulation) and NM (Negative Modulation)

Lu, et al. [7] had analyzed the behaviors of transformed coefficients under attacks. In principle, there are four possible types of modulations: Modu(+, +), Modu(+, −), Modu(−, +), and Modu(−, −), where Modu(+/−, −/+) denotes a positive/negative transformed coefficient modulated with a negative/positive watermark quantity. No matter whether the DCT or the wavelet domain is employed, the probabilities of occurrence of the four types of modulations are all very close to 0.25.

They further classified the behaviors of attacks into two categories. The first category contains those attacks like compression and blurring, which tend to decrease the magnitudes of most of the transformed coefficients. Under these circumstances, it is hoped that every transformed coefficient can be modulated with a quantity that has different sign. The reason is that it can adapt to compression-style attacks and enables more than 50% of the modulated targets to contribute a bigger positive value to the detector response. Only Modu(+, −) and Modu(−, +) will contribute positively to the detector response. The second category contains those attacks such as sharpening and histogram equalization, which have the tendency of increasing most of the magnitudes of transformed coefficients. Only Modu(+, +) and Modu(−, −) will contribute positively to the detector response. Lu, et al. emphasized that the random modulation strategy does not help the detector response.

Scenarios in the attacking process are illustrated in Fig. 2. No matter whether the positive modulation or the negative modulation is employed, the modulated wavelet coefficient can effectively resist the attack in scenario 1 and 2. However, the modulated wavelet coefficient is unable to resist the attack alone in scenario 3 if the strength of the attack is larger than that of the modulation. If a watermarking algorithm simultaneously employs the positive modulation and the negative modulation in embedding a watermark bit, it can succeed in resisting the attack in scenario 3 (as cocktail watermarking did in [7], which simultaneously embedded two watermarks in complementary roles).

Since the DEW scheme and the WTQ scheme only employed the philosophy of negative modulation, the detector was unable to bring the brilliant results under any kind of attacks mentioned above. Moreover, the scheme can be easily defeated by the attacker if it only employs unilateral modulation, regardless of the positive modulation or the negative modulation. Thus, a good differential energy watermarking algorithm should take both modulated methods into account for higher detector response and better security.

4.2 Wavelet Tree Structure

We employ the same wavelet tree structure as depicted in the WTQ scheme. However, each tree can be extended to involve high-frequency components as illustrated in Fig. 4.

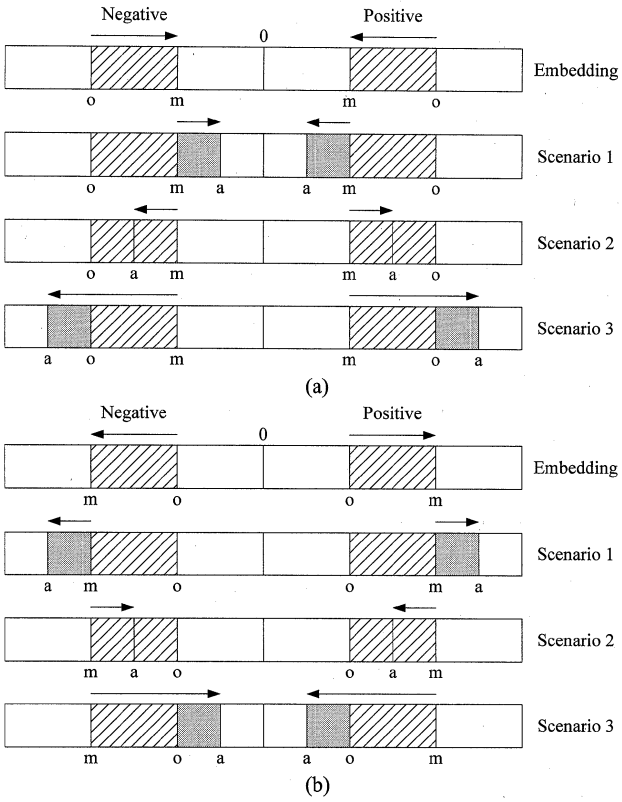


Fig. 2 Scenarios in the attacking process. (a) Positive modulation. (b) Negative modulation. "o" denotes the original wavelet coefficient, "m" indicates the modulated wavelet coefficient, and "a" means the attacked wavelet coefficient.

Suppose that a 512×512 image is transformed, each tree will be a collection of 85 wavelet coefficients, one coefficient from level 4, 4 coefficients from level 3, 16 coefficients from level 2, and 64 coefficients from level 1. There will be two parameter set S_1 and S_2 for WTGM [8] and they will be discussed in Sect. 5.

4.3 Super Trees Selection for WTGM

The idea of sum-of-subsets [5] for selecting supertrees is adopted in WTGM to securely embed the watermark. The sum-of-subsets problem can be formulated as following: There are n positive integers (weights) w_i and a positive integer W. The goal is to find all subsets of the integers that sum to W. For example, $S = \{11, 13, 24, 7\}$, $W = 31$, then there are two subsets: $\{11, 13, 7\}$, $\{24, 7\}$.

In fact, the sum-of-subsets problem itself is an NP-complete problem. Das, Maitra and Mitra [5] used this method to resolve the vulnerability of the DEW scheme. As a key factor, it renders the idea how the coefficients are grouped together with a closed value of energy aggregation. If we treat the energy of a tree as an element in the set S, and if we can find out which trees can be grouped together to form the so-called super tree, we can use this principle to modulate the coefficients according the watermark bit embedded.

Suppose that each watermark bit is embedded using one super tree, half of a super tree is used for PM and the other is used for NM. We use the term super tree to refer to the collection of n trees (i.e. 1 super tree = n trees). A particular super tree can be divided into two sub-super trees, each containing $n/2$ trees. The energy of a tree t is defined as the sum of absolute values of $q - p+1$ wavelet coefficients. The energies of sub-super tree A and sub-super tree B are given by:

$$E_A(p,q,n) = \sum_{t=0}^{\frac{n}{2}-1} \sum_{i=p}^{q} |\theta_{i,t}| \tag{1}$$

$$E_B(p,q,n) = \sum_{t=\frac{n}{2}}^{n-1} \sum_{i=p}^{q} |\theta_{i,t}| \tag{2}$$

where $\theta_{i,t}$ denotes the ith wavelet coefficient in the tree t, p and q denotes the coefficient number used to do the modulation from p to q ($0 \le p \le 84$, $0 \le q \le 84$). Any two sub-super trees with $E_A \approx E_B$, i.e. $|E_A - E_B| \le \delta$, will be suitable for modulation. $|E_A - E_B| \le \delta$ is just a criterion judging whether the energy of sub-super tree A and that of B differ by less than or equal to some small quantity δ.

4.4 CSF (Contrast Sensitive Function) of HVS

For watermarked images, there has been a need for good metrics for image quality that incorporates properties of the HVS. The visibility thresholds of visual signals are studied by psychovisual measurements to determine the thresholds. These measurements were performed on sinusoidal gratings with various spatial frequencies and orientations by given viewing conditions. The purpose of such study was to determine the contrast thresholds of gratings by the given frequency and orientation. Contrast as a measure of relative variation of luminance for periodic pattern such as a sinusoidal grating is given by the equation

$$C = (L_{\max} - L_{\min})/(L_{\max} + L_{\min}) \tag{3}$$

where $L_{\max}$ and $L_{\min}$ are maximal and minimal luminance of a grating. Reciprocal values of contrast thresholds express the contrast sensitivity (CS), and Mannos and Sakrison [9] originally presented a model of the contrast sensitive function (CSF) for luminance (or grayscale) images is given as follows:

$$H(f) = 2.6 * (0.0192 + 0.114 * f) * e^{-(0.114*f)^{1.1}} \tag{4}$$

where $f = \sqrt{f_x^2 + f_y^2}$ is the spatial frequency in cycles/degree of visual angle (f_x and f_y are the spatial frequencies in the horizontal and vertical directions, respectively). Figure 3 depicts the CSF curve which characterizes luminance sensitivity of the HVS as a function of normalized spatial frequency. According to the CSF curve, we can see that the HVS is most sensitive to normalized spatial frequencies between 0.025 and 0.125 and less sensitive to low and high frequencies [10]. Therefore, this knowledge from CSF

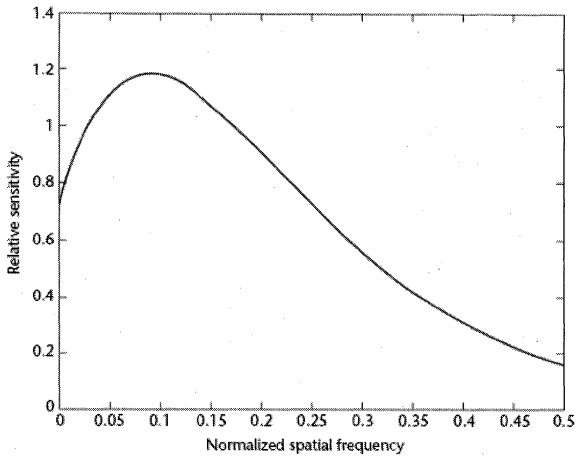


Fig. 3 Luminance CSF (Courtesy of [10]).

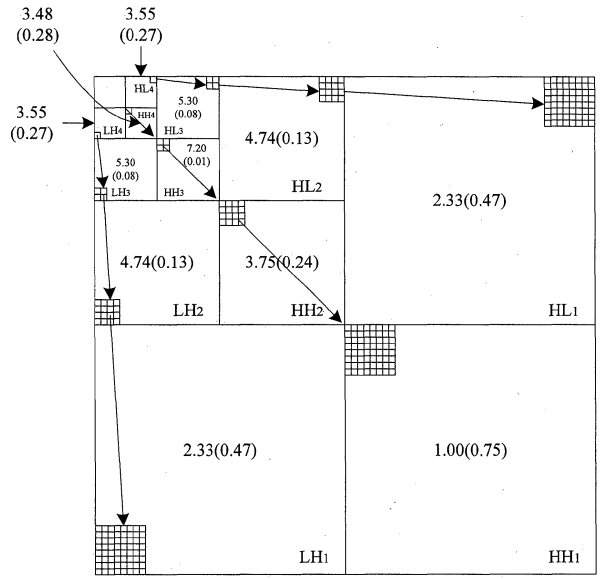


Fig. 4 A four-level wavelet tree structure. The coefficients correspond to the same spatial location are grouped together. Each tree consists of one coefficient from level 4, 4 coefficients from level 3, 16 coefficients from level 2, and 64 coefficients from level 1. $r^k(\beta^k)$ values for each level k are indicated at the center of each band.

can be used to develop a simple image independent HVS model.

CSF masking [10]–[13] is one way to apply the CSF in the discrete wavelet domain. CSF masking refers to the method of weighting the wavelet coefficients according to their perceptual importance. Some well-designed CSF masks which transforms the CSF curve in Fig. 3 into perceptual importance weight are presented in [11]. Huang and Tang [10] use the same method led to 12-weight DWT CSF mask in the five-level wavelet transform. Figure 4 illustrates the 12-weight DWT CSF mask with the weights shown for each subband.

We use the square function in [10] to approximate the effect of CSF masking. The adequate modulation rate β^k for each subband is determined by:

$$\beta^k = 0.01 + \frac{(7.20 - r^k)^2}{7.20^2} \tag{5}$$

where k denotes the decomposed level and r^k represents the wavelet coefficient CSF of the perceptual importance weight as Fig. 4 shows. The level 1 has the largest rate for modulation, which corresponds to high-frequency components. The level 3 has the smallest rate for modulation.

Under the circumstances the sum-of-subsets is employed, the actual modulation quantity of low-frequency components will be relatively small since they have larger energies. Contrarily, the actual modulation quantity of high-frequency components will be relatively large since they have smaller energies. In our study, low-frequency components can tolerate more common signal processing while high-frequency components can tolerate more geometric attacks. The usage of high-frequency components is pretty different from the WTQ scheme for its nature of watermarking.

4.5 NVF (Noise Visibility Function) of HVS

S. Voloshynovskiy et al. [14] presented a stochastic approach based on the computation of a NVF (Noise Visibility Function) that characterizes the local image properties and identifies texture and edge regions. This allows us to determine the optimal watermark locations and strength for the watermark embedding stage.

Their argument: the channel capacity is not uniform, i.e. the noise is more visible in flat areas and less visible in regions with edges and textures. Accordingly, when the local variance is small, the image is flat, and a large enough variance indicates the presence of edges or highly texture areas. The adaptive scheme based on NVF calculated from stationary GG model is the best model in our simulation, which is defined as follows:

$$NVF(x,y) = \frac{w(x,y)}{w(x,y) + \sigma_I^2} \tag{6}$$

where $w(x,y) = \gamma[\eta(\gamma)]^\gamma / \|r(x,y)\|^{2-\gamma}$ and σ_I^2 is the global variance. $\eta(\gamma) = \sqrt{\Gamma(3/\gamma)/\Gamma(1/\gamma)}$, $\Gamma(s) = \int_0^\infty e^{-u}u^{s-1}du$ (gamma function) and $r(x,y) = (I(x,y) - \bar{I}(x,y))/\sigma_I$. γ is the shape parameter and $r(x,y)$ is determined by the local mean and the local variance. For most of real images, the shape parameter is in the range $0.3 \le \gamma \le 1$.

Even the PSNR value is very close, the image quality is quite different. The combination of the CSF and NVF can effectively decrease the visibility and enhance the energy of the watermark. Since the CSF constraints the modulation rate which retains a better visual effect, while at the same time the NVF enhance the watermark strength in texture and edge regions which renders a better detector response.

4.6 WTGM Algorithm

We summarize the ideas mentioned above in the following algorithms, which integrate the advantages of wavelet

tree structure, sum-of-subsets for supertree selection, positive/negative modulation for watermark embedding and the CSF and NVF of the HVS into the WTGM. To quantify the existence of the watermark, the normalized correlation coefficient (NC) will be examined in order to identify the existence of the watermark. The formula of normalized correlation coefficient is as follows:

$$\rho(W, W') = \frac{\sum_{m=1}^{N_w} w_m w'_m}{N_w} \tag{7}$$

The coefficient value is within -1 and 1.

The complete design of the proposed algorithm is summarized as following:

WTGM Watermark Embedding:

1) Generate a seed by mapping a signature/text through a one-way deterministic function. Obtain a PN sequence W of length N_w using the seed.
2) Compute wavelet coefficients of a host image. Group the coefficients to form trees.
3) Randomly arrange the trees using some pseudorandom generator and group them in various super trees. Each super tree should be divided in two sub-super trees such that $E_A \approx E_B$. Store this group information which we call the image key K.
4) For each watermark bit w_i ($i = 0$ to $N_w - 1$) do
   a) Select the ith super tree consisting of n trees.
   b) Choose α.
   c) If ($w_i = -1$) then
      i) $\theta_{i,t} = \theta_{i,t} * (1 + \alpha * \beta^k * \gamma^k_{x,y})$ for $t = 0, \ldots, (n/2) - 1$, and $i = p, \ldots, q$. (PM for sub-super tree A)
      ii) $\theta_{i,t} = \theta_{i,t} * (1 - \alpha * \beta^k * \gamma^k_{x,y})$ for $t = (n/2), \ldots, n - 1$, and $i = p, \ldots, q$. (NM for sub-super tree B)
   d) else
      i) $\theta_{i,t} = \theta_{i,t} * (1 - \alpha * \beta^k * \gamma^k_{x,y})$ for $t = 0, \ldots, (n/2) - 1$, and $i = p, \ldots, q$. (NM for sub-super tree A)
      ii) $\theta_{i,t} = \theta_{i,t} * (1 + \alpha * \beta^k * \gamma^k_{x,y})$ for $t = (n/2), \ldots, n - 1$, and $i = p, \ldots, q$. (PM for sub-super tree B)
5) Arrange back the modulated trees to their original positions.
6) Pass the modified wavelet coefficients through the inverse DWT to obtain a watermarked image.

Note:

1) The watermark W is a binary PN sequence of ± 1.
2) The length of the watermark N_w = the number of super trees. The max value of $N_w = 1536$ for a 512×512 image under 4 level wavelet decomposition. Since each watermark bit is embedded using one super tree where half of a super tree is used for PM and the other is used for NM, WTGM needs at least 12 bits to mark each super tree ($2^{12} = 4096 \geq 1536 \times 2$). Therefore, the image key K needs $12 \times 2 \times N_w$ bits for recording the ordering information under the sum-of-subsets principle. If N_w equals to the value of 1536, the image key K will be $12 \times 2 \times 1536$ bits which is 4608 bytes without any compression. The amount of image key K can be further reduced by Huffman coding or zip compression tool and the data size would be comparatively small.
3) Since the image key K is the important and secret information, the secure data transmission is a critical issue for watermark extraction. The research of cryptography studies the advanced algorithms of encryption and decryption to guarantee confidentiality and authenticity of data is needed for secure communication. Data encryption algorithms which systematically change the contents of the secret data can efficiently encode the data so they can't be read by anyone who doesn't have the proper key to unscramble them. Interest readers could refer the cryptography study [15] for more detailed information for application. For practical use, there are free tools available to encrypt and decrypt files like PGP [16] which is a public-key encryptions system that lets individuals secure their documents with extremely strong encryption algorithms and long keys.
4) α originally denotes the fractional change required to enforce the required energy difference, i.e., after the modification, we need $|(E'_A - E'_B)/(E_A + E_B)| \geq \alpha$ (E'_A and E'_B are E_A, E_B after the modification). If the HVS is employed, it stands for the strength of the watermark.
5) β^k is the CSF embedding parameter and $\gamma^k_{x,y} = 1 - NVF(x, y)$ is the NVF embedding parameter where the formula (5) is used.
6) If $\beta^k = 1$ and $\gamma^k_{x,y} = 1$, the HVS is not employed.
   If $\beta^k = 1$, the CSF is not employed.
   If $\gamma^k_{x,y} = 1$, the NVF is not employed.
7) The information of super tree list will be stored during the embedding procedure but the random number generator algorithm is not necessary to be recorded. Most pseudo-random generator algorithms produce sequences which are uniformly distributed [17] and common classes of these algorithms are linear congruential generators, lagged Fibonacci generators, linear feedback shift registers, generalized feedback shift registers, Blum Blum Shub, Fortuna, and the Mersenne twister. Therefore, it is very flexible for WTGM to apply any good random number generator to arrange the wavelet trees into various super trees under the principle of sum-of-subsets.

WTGM Watermark Extraction

1) Generate a seed by mapping a signature/text through a one-way deterministic function. Obtain a PN sequence W of length N_w using the seed.
2) Compute wavelet coefficients of a host image. Group the coefficients to form trees.
3) Reorganize the trees using the image key K.
4) For each watermark bit w_i ($i = 0$ to $N_w - 1$) do
   a) Select the ith super tree consisting of n trees.
   b) Calculate E_A and E_B.
   c) If ($E_A > E_B$) then $w_i = -1$.

else $w_i = 1$.

5) Compute the normalized correlation ρ.
6) If ρ is above the threshold ρ_T, the watermark W exists; otherwise, it does not exist.

5. Experiment Results

To evaluate the performance of the proposed method, the 512 × 512 Lena, Goldhill and Peppers images with 8 bits/pixel resolution are used for watermarking. We employ a four-level wavelet transform and a watermark sequence of length 512. Therefore, a super tree consists of 6 trees, half the trees are used for PM and the others are used for NM.

The experiments are divided into two parts. The first part WTGM(S_1) (Watermarking Parameter Set 1 (S_1)) uses coefficient number 1–21 (i.e. $p = 0$, $q = 20$) corresponds to relatively low-frequency components (level 2, 3 and 4 of DWT) for watermarking, which is the same as the WTQ scheme. The second part WTGM(S_2) uses coefficient number 6–85 (i.e. $p = 5$, $q = 84$) corresponds to relatively high-frequency components (level 1 and 2 of DWT).

In order to make the fair comparison, all the watermarked images will be set at the same PSNR values shown as in [4] of WTQ algorithm since it is the typically representative wavelet tree based approach. To compare with the WTQ scheme, we set the value of α (watermark strength) to meet the PSNR values of 38.2, 38.7 and 39.8 dB for Lena, Goldhill and Peppers since the setting is the same as in WTQ, respectively, as shown in Table 1. With watermark length $N_w = 512$, the threshold ρ_T of NC is chosen to be 0.23 for a false positive probability of 1.03×10^{-7}. The wavelet filters used in this study for the wavelet tree watermarking is the CDF 9/7 filters which are also used in WTQ.

5.1 Visual Quality Comparison

From Fig. 5, the two different parameter setting will result in different image quality while even the PSNR is kept at the same. Compared by Figs. 5(b) and 5(d), the error images demonstrate the errors between the watermarked images and the original by setting (S_1) and (S_2) are different. WTGM(S_2) watermarked image will have more high frequency signals than WTGM(S_1) watermarked image.

Figure 6 shows the visual quality under different watermarking parameter settings. The watermarked Lena images are all with PSNR values of 38.2 dB. Figure 6(b) uses coefficient number 1–21 to embed the watermark, which corresponds to the subbands used in the WTQ scheme. Figure 6(c) uses coefficient number 6–85 to embed the watermark. Without consideration to the HVS, even the modulation rate is small ($\alpha = 0.177$ and 0.378), there are obvious artifacts in the region marked with a rectangle. The employment of the HVS apparently improves the visual quality of the watermarked image, even it has larger modulation rate (as shown in Fig. 6(d)).

Figure 7 is another example of why the HVS is important. We intensify the strength of watermark so that the

Table 1 The parameter settings for Lena, Goldhill and Peppers images.

WTGM	Image – Watermark Strength (α)		
	Lena	Goldhill	Peppers
S_1	0.968	0.959	0.706
S_2	2.131	1.346	1.358
PSNR (dB)	38.2	38.7	39.8

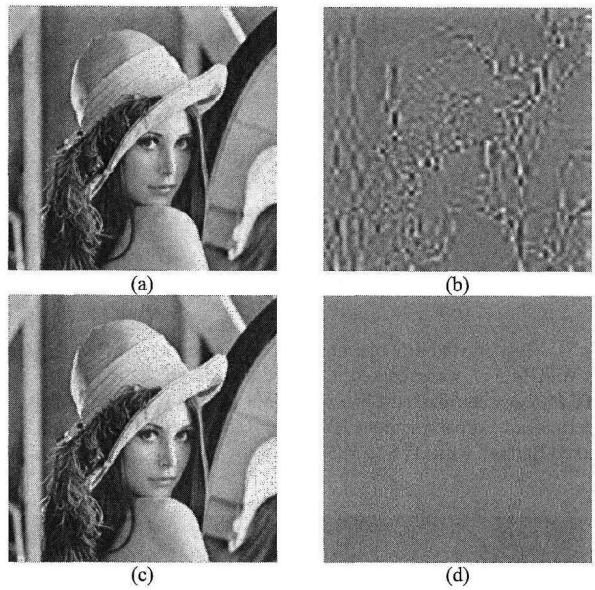


Fig. 5 Watermarked images and error images. (a) WTGM(S_1) watermarked Lena image at PSNR=38.2 dB. (b) Scaled error image between (a) and the original image. (c) WTGM(S_2) watermarked Lena image at PSNR=38.2 dB. (d) Scaled error image between (c) and the original image.

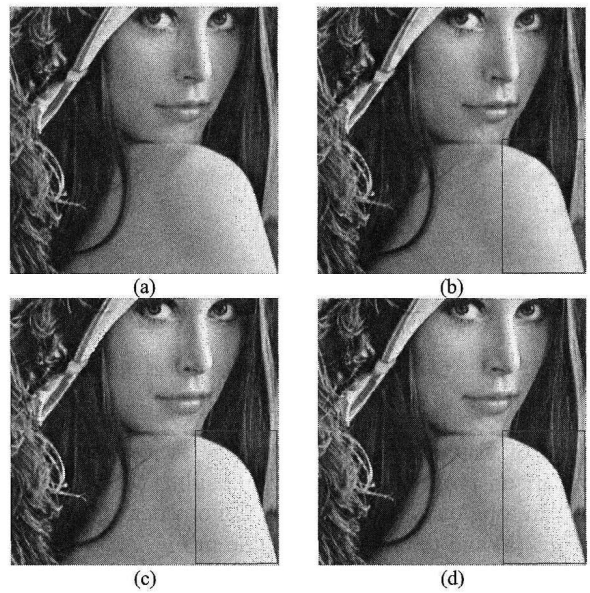


Fig. 6 Close-ups for comparison with visual effects (PSNR = 38.2 dB). (a) Original image. (b) WTGM(S_1) watermarked Lena without HVS (α = 0.177). (c) WTGM(S_2) watermarked Lena without HVS (α = 0.378). (d) WTGM(S_2) watermarked with HVS (α = 2.131).

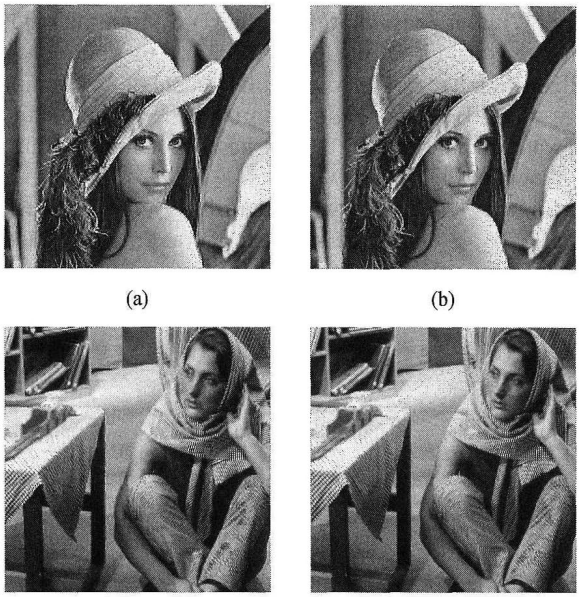


Fig. 7 Test on visibility of embedding of watermark (PSNR = 25.0 dB). (a) WTGM(S_2) watermarked Lena without HVS (α = 1.731). (b) WTGM(S_2) watermarked Lena with HVS (α = 9.741). (c) WTGM(S_2) watermarked Barbara without HVS (α = 0.872). (d) WTGM(S_2) watermarked Barbara with HVS (α = 3.374).

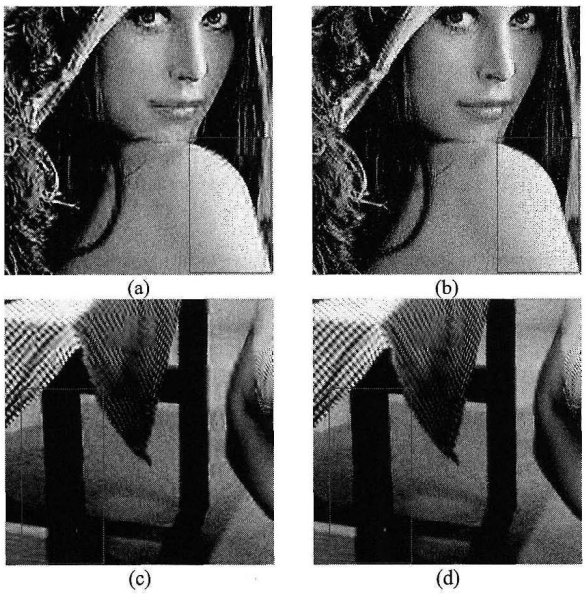


Fig. 8 Close-ups for comparison with Figs. 7(a)–(d).

quality of watermarked image will be as low as 25 dB for comparison purpose. From these results, we can see that there are obvious artifacts in the regions near the shoulder in Fig. 7(a) and the foot of the table in Fig. 7(c). The HVS can effectively decrease the visibility of the watermark (as shown in Figs. 7(b) and (d)). Figures 8(a)–(d) are the close-ups of the images in Figs. 7(a)–(d).

Table 2 Watermarks extracted from JPEG compressed watermarked images.

JPEG (QF)	Image								
	Lena			Goldhill			Peppers		
	WTGM		WTQ	WTGM		WTQ	WTGM		WTQ
	S_1	S_2		S_1	S_2		S_1	S_2	
100	1.000	1.000		1.000	1.000		1.000	1.000	
90	1.000	1.000	1.00	1.000	1.000	1.00	1.000	1.000	1.00
80	1.000	1.000		1.000	1.000		0.996	0.988	
70	0.996	1.000	0.57	1.000	0.996	0.93	0.988	0.809	0.97
60	0.977	0.980		1.000	0.996		0.969	0.719	
50	0.977	0.941	0.26	0.996	0.988	0.71	0.938	0.629	0.70
40	0.961	0.859	0.23	0.984	0.973	0.52	0.898	0.598	0.54
30	0.945	0.770	0.15	0.953	0.922	0.23	0.832	0.488	0.34
20	0.828	0.617		0.906	0.809		0.699	0.363	
10	0.602	0.355		0.664	0.430		0.469	0.180	
0	0.102	-0.012		0.002	0.051		0.094	0.008	

5.2 Common Image Processing Attacks

1) JPEG Compression Attacks

In this experiment, we perform JPEG compression with different quality factors (QF) on the watermarked image. The extracted results and NC values are depicted in Table 2. From these results, we can see that the proposed algorithm is robust to JPEG compression. For all cases, the extracted watermarks are with relatively high-NC values. The result of WTGM(S_1) is superior to that of WTGM(S_2). Even for the case that QF is equal to 20, we can still detect the embedded watermark.

Since the setting for S_2 reserves the watermark in the level 1 component, JPEG intentionally removes the high frequency components which make setting S_1 perform better than setting S_2. Therefore, the results from Table 2 are reasonable.

2) SPIHT Compression Attacks

SPIHT (Set Partitioning in Hierarchical Trees) is an image compression algorithm that exploits the inherent similarities across subbands in a wavelet decomposition of an image. It implies uniform quantization and bit allocation applied after wavelet decomposition. Table 3 shows the extracted NC values and corresponding PSNR values between original image and attacked image. From these results, we can see that the proposed algorithm can tolerate the incidental distortions induced by high-quality SPIHT compression. Since SPIHT first removes the high frequency components during the rate reduction, the results of WTGM(S_1) is also superior to those of WTGM(S_2).

3) JPEG2000 Compression Attacks

JPEG2000 [18] is a new image compression standard which has good performance in high bit rate coding. It adopts wavelet transform instead of discrete cosine transform to utilize the intersubband correlation. Table 4 shows the extracted NC values and corresponding PSNR values between original image and attacked image. Since there is no data from WTQ results under JPEG2000 attack, the results under SPIHT attack are shown for comparison purpose. From these results, we can see that the proposed WTGM al-

Table 3 Watermarks extracted from SPIHT compressed watermarked images. (a) Lena. (b) Goldhill. (c) Peppers.

Bit rate	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ NC	WTQ PSNR
0.7	1.000	35.430	1.000	35.942	0.85	36.7
0.6	0.996	35.084	0.992	35.804	0.83	35.2
0.5	0.996	34.660	0.988	35.269	0.85	34.6
0.4	0.980	34.089	0.988	34.645	0.41	34.3
0.3	0.957	33.223	0.770	33.985	0.21	33.1

(a)

Bit rate	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ NC	WTQ PSNR
0.7	1.000	33.171	1.000	33.403	0.35	34.1
0.6	1.000	32.655	1.000	32.870	0.27	33.2
0.5	0.992	32.046	0.973	32.318	0.23	32.9
0.4	0.969	31.287	0.898	31.662	0.02	32.4
0.3	0.945	30.425	0.875	30.750	-0.06	31.7

(b)

Bit rate	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ NC	WTQ PSNR
0.7	0.984	35.082	1.000	35.553	0.85	34.9
0.6	0.980	34.776	0.996	35.227	0.71	34.7
0.5	0.988	34.391	0.980	34.783	0.65	34.4
0.4	0.941	33.923	0.980	34.041	0.66	33.6
0.3	0.926	33.135	0.641	33.721	0.36	33.1

(c)

Table 4 Watermarks extracted from JPEG2000 compressed watermarked images. (a) Lena. (b) Goldhill. (c) Peppers.

Bit rate	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ(SPIHT) NC	WTQ(SPIHT) PSNR
0.7	0.996	34.966	0.984	35.268	0.85	36.7
0.6	0.984	34.552	0.953	34.979	0.83	35.2
0.5	0.973	34.101	0.926	34.551	0.85	34.6
0.4	0.949	33.513	0.844	34.001	0.41	34.3
0.3	0.934	32.575	0.715	33.162	0.21	33.1

(a)

Bit rate	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ(SPIHT) NC	WTQ(SPIHT) PSNR
0.7	0.992	32.886	0.996	32.984	0.35	34.1
0.6	0.988	32.360	0.980	32.461	0.27	33.2
0.5	0.945	31.721	0.895	31.981	0.23	32.9
0.4	0.922	30.952	0.863	31.221	0.02	32.4
0.3	0.941	30.096	0.734	30.382	-0.06	31.7

(b)

Bit rate	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ(SPIHT) NC	WTQ(SPIHT) PSNR
0.7	0.945	34.638	0.996	34.575	0.85	34.9
0.6	0.926	34.334	0.992	34.173	0.71	34.7
0.5	0.930	33.997	0.949	34.050	0.65	34.4
0.4	0.883	33.511	0.813	33.606	0.66	33.6
0.3	0.766	32.737	0.621	33.010	0.36	33.1

(c)

gorithm can tolerate the incidental distortions induced by JPEG2000 compression. Since JPEG2000 first removes the high frequency components during the rate reduction, the results of WTGM(S_1) is also superior to those of WTGM(S_2) which has similar performance as shown in Table 3.

Table 5 Watermarks extracted from spatial-domain-attacked watermarked images. (a) Lena. (b) Goldhill. (c) Peppers.

Operations	WTGM S_1	WTGM S_2	WTQ
Histogram Equalization	0.824	1.000	
Image Cropping (25% upper-left corner)	0.574	0.840	
Brightness Enhancement (10%)	1.000	1.000	
Contrast Enhancement (10%)	1.000	1.000	
Median Filtering (2×2)	0.797	0.961	0.38
Median Filtering (3×3)	0.891	0.867	0.51
Median Filtering (4×4)	0.555	0.574	0.23
Gaussian Filtering	0.457	0.680	0.64
Sharpening	0.629	1.000	0.46
Rescale (50%)	0.918	0.578	
Rescale (90%)	0.859	0.996	
Rescale (150%)	0.918	1.000	

(a)

Operations	WTGM S_1	WTGM S_2	WTQ
Histogram Equalization	0.707	0.996	
Image Cropping (25% upper-left corner)	0.395	0.703	
Brightness Enhancement (10%)	1.000	1.000	
Contrast Enhancement (10%)	1.000	1.000	
Median Filtering (2×2)	0.781	0.926	0.35
Median Filtering (3×3)	0.922	0.848	0.56
Median Filtering (4×4)	0.648	0.520	0.24
Gaussian Filtering	0.578	0.801	0.56
Sharpening	0.793	1.000	0.39
Rescale (50%)	0.934	0.652	
Rescale (90%)	0.828	1.000	
Rescale (150%)	0.930	0.996	

(b)

Operations	WTGM S_1	WTGM S_2	WTQ
Histogram Equalization	0.711	1.000	
Image Cropping (25% upper-left corner)	0.375	0.809	
Brightness Enhancement (10%)	1.000	1.000	
Contrast Enhancement (10%)	1.000	1.000	
Median Filtering (2×2)	0.707	0.813	0.46
Median Filtering (3×3)	0.926	0.637	0.71
Median Filtering (4×4)	0.559	0.398	0.25
Gaussian Filtering	0.355	0.418	0.74
Sharpening	0.551	0.996	0.62
Rescale (50%)	0.914	0.336	
Rescale (90%)	0.660	0.941	
Rescale (150%)	0.828	0.922	

(c)

4) Spatial-Domain Image Processing Attacks

Several spatial-domain image processing techniques, including histogram equalization, image cropping, brightness enhancement, contrast enhancement, median filtering, Gaussian filtering, sharpening, and rescale are performed on the watermarked image. The extracted results are depicted in Table 5. For all cases, the watermark information therein can be successfully recognized. Especially for those cases of histogram equalization, Gaussian filtering and sharpening, the result of WTGM(S_2) is superior to that of WTGM(S_1). Except for the case of Peppers Gaussian filtered image, the proposed algorithm can outperform the

Table 6 Watermarks extracted from shifted watermarked images.

Pixel Shift	Image								
	Lena			Goldhill			Peppers		
	WTGM		WTQ	WTGM		WTQ	WTGM		WTQ
	S_1	S_2		S_1	S_2		S_1	S_2	
4	0.246	0.961		0.477	0.980		0.270	0.926	
5	0.133	0.867	0.28	0.309	0.922	0.36	0.160	0.867	0.32
6	0.137	0.891	0.34	0.277	0.902	0.35	0.184	0.859	0.34
7	0.145	0.730	0.29	0.250	0.797	0.41	0.117	0.766	0.29
8	0.180	0.785	0.81	0.227	0.793	0.84	0.164	0.746	0.92
9	0.145	0.547	0.26	0.156	0.637	0.29	0.145	0.543	0.29
10	0.129	0.531	0.19	0.160	0.617	0.21	0.164	0.563	0.26
11	0.125	0.363		0.113	0.504		0.160	0.395	
12	0.086	0.367		0.129	0.406		0.129	0.359	
13	0.023	0.199		0.066	0.285		0.078	0.227	

Table 7 Watermarks extracted from rotated watermarked images.

Rotation	Image								
	Lena			Goldhill			Peppers		
	WTGM		WTQ	WTGM		WTQ	WTGM		WTQ
	S_1	S_2		S_1	S_2		S_1	S_2	
-0.25	0.805	0.988	0.32	0.727	0.984	0.38	0.590	0.922	0.39
-0.50	0.570	0.980	0.23	0.559	0.949	0.27	0.289	0.852	0.25
-0.75	0.402	0.953	0.24	0.434	0.887	0.25	0.270	0.805	0.25
-1.00	0.250	0.887	0.16	0.363	0.824	0.14	0.242	0.727	0.16
-1.50	0.195	0.711		0.184	0.703		0.199	0.559	
-2.00	0.133	0.555		0.129	0.531		0.156	0.473	
-2.50	0.117	0.406		0.004	0.367		0.113	0.266	
-3.00	0.098	0.227		-0.039	0.238		0.051	0.203	
0.25	0.742	0.973	0.37	0.770	0.984	0.33	0.578	0.914	0.41
0.50	0.469	0.957	0.29	0.555	0.953	0.24	0.309	0.852	0.30
0.75	0.328	0.902	0.26	0.465	0.910	0.21	0.289	0.801	0.26
1.00	0.215	0.875	0.24	0.359	0.875	0.15	0.184	0.785	0.17
1.50	0.141	0.719		0.293	0.719		0.117	0.645	
2.00	0.133	0.523		0.164	0.527		0.078	0.473	
2.50	0.102	0.414		0.145	0.359		0.090	0.324	
3.00	0.047	0.332		0.070	0.336		0.051	0.254	

WTQ scheme with relatively high-NC values.

5.3 Geometric Attacks

1) Pixel Shifting Attacks (Circular Shift)

This kind of attacks is done by shifting the pixels circularly. Here, we shift the pixels to the left. Apparently WTGM(S_1) is unable to resist such attacks as shown in Table 6. Contrarily, WTGM(S_2) can resist a shift of up to 12 pixels for Lena and Peppers images and 13 pixels for Goldhill image. For the former has lower modulation rate than that of the latter.

2) Rotation Attacks (Rotation and Scaling)

The attack is done by rotating the image by a small angle, scaling the rotated image, and cropping the scaled image to the original image size. StirMark [19] software is adopted here for this attack since it provides the described testing functions. This rotation and scaling is a geometrical attack in the spatial domain. We rotate the watermarked image from 0.25° to 3° in clockwise and counter-clockwise directions. The extracted results are shown in Table 7. From these results, we can see that the WTGM(S_2) can resist a rotation of up to 3° for Goldhill image and 2.5° for Lena and Peppers images.

Table 8 Watermarks extracted from multiple watermarked images. (a) Lena. (b) Goldhill. (c) Peppers.

Number of Watermarks	WTGM				WTQ	
	S_1		S_2			
	NC	PSNR	NC	PSNR	NC	PSNR
1	0.965	35.129	0.965	34.181	0.65	35.50
2	0.762	33.885	0.855	31.152	0.41	32.78
3	0.734	32.606	0.727	28.778	0.27	29.35
4	0.617	31.520	0.590	26.807	0.11	28.05
6	0.523	29.673	0.395	24.083		

(a)

Number of Watermarks	WTGM				WTQ	
	S_1		S_2			
	NC	PSNR	NC	PSNR	NC	PSNR
1	0.992	35.348	0.980	35.223	0.79	35.26
2	0.863	33.736	0.836	32.962	0.45	31.50
3	0.766	32.345	0.766	31.251	0.31	29.71
4	0.738	31.242	0.711	29.869	0.18	28.57
6	0.621	29.811	0.555	27.838		

(b)

Number of Watermarks	WTGM				WTQ	
	S_1		S_2			
	NC	PSNR	NC	PSNR	NC	PSNR
1	0.977	36.753	0.973	36.110	0.80	34.53
2	0.844	34.713	0.844	33.456	0.53	31.99
3	0.734	33.326	0.730	31.360	0.31	30.19
4	0.754	32.446	0.637	29.958	0.22	28.81
6	0.609	30.770	0.469	28.111		

(c)

5.4 Security Measurement

1) Multiple Watermarking

For algorithms well-known to all, the attacker may apply one or more watermarks using the same wavelet tree group modulation technique in an attempt to disturb the detector or to destroy the embedded watermark. Table 8 shows the results of the watermarked images attacked through multiple watermarking. From these results, we can see that even the PSNR value of attacked image is fallen into 25 dB, the watermark still can be detected.

2) Bitplane Removal

Bitplane removal is one of the major strategies used to defeat the WTQ scheme. We perform this attack designated on the embedded subbands, which reduces the impact on watermarked images. Table 9 shows that the proposed algorithm can resist 7 and 8 bitplanes removed for WTGM(S_1) and WTGM(S_2), respectively. Under which the PSNR values of attacked images are fallen into 26 and 29 dB.

5.5 Complexity of WTGM with Human Vision System

The computation complexity of WTGM with Human Vision System is also low from the view of mathematical analysis. The whole complexity should be discussed for wavelet transform, sum-of-subsets, CSF and NVF calculation respectively.

Suppose the synthesis filters are h (low-pass) and g (high-pass) for wavelet transform. Take $|h| = 2N$, $|g| = 2M$, and assume $M \geq N$. The cost of the standard algorithm for CDF 9/7 filters is $4(N + M)+2$ and could be speeded up by

Table 9 Watermarks extracted from bitplane-removed watermarked images. (a) Lena. (b) Goldhill. (c) Peppers.

Number of Bitplanes Removed	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ NC	WTQ PSNR
1	1.000	38.120	1.000	38.252	1.00	36.81
2	1.000	37.852	1.000	38.106	1.00	34.72
3	1.000	37.043	1.000	37.247	0.99	30.41
4	0.980	35.423	0.992	35.583	0.52	24.28
5	0.781	32.736	0.879	33.524	0.11	18.47
6	0.480	29.619	0.891	31.500		
7	0.340	26.715	0.895	30.139		
8	0.164	24.442	0.887	29.814		

(a)

Number of Bitplanes Removed	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ NC	WTQ PSNR
1	1.000	38.599	1.000	38.489	1.00	36.07
2	1.000	38.277	1.000	37.779	1.00	33.72
3	1.000	37.192	1.000	36.144	0.97	28.87
4	0.996	34.832	1.000	33.675	0.38	22.72
5	0.930	31.699	0.875	31.211	0.14	16.18
6	0.652	28.666	0.523	29.397		
7	0.402	26.190	0.426	28.365		
8	0.133	24.574	0.395	28.281		

(b)

Number of Bitplanes Removed	WTGM S_1 NC	WTGM S_1 PSNR	WTGM S_2 NC	WTGM S_2 PSNR	WTQ NC	WTQ PSNR
1	1.000	39.698	1.000	39.544	0.99	35.97
2	1.000	39.334	1.000	38.785	0.96	33.64
3	1.000	38.210	1.000	37.359	0.90	28.76
4	0.949	36.181	0.980	35.389	0.64	22.71
5	0.734	33.424	0.594	33.394	0.14	16.93
6	0.543	30.236	0.457	31.511		
7	0.340	26.987	0.379	29.728		
8	0.246	24.418	0.348	29.179		

(c)

the lifting algorithm in [20] to $2(N + M + 2)$. The computation of wavelet transform is linear time mathematics.

The sum-of-subsets problem itself is a known NP-complete problem [5]. However, WTGM is not dealing with a real sum-of-subsets problem but a sum-of-subsets idea instead. Empirical study shows that the implementation of sum-of-subsets in WTGM can actually be applied by quicksort [21] to order the supertrees based on the tree energy to get such an arrangement easily. Therefore, its time complexity requires only about $(2 + 2\ln 2)R$ comparisons if R items are sorted and the complexity of the quicksort-based selection is linear-time on the average [21].

On the other hand, CSF masking is employed to apply the CSF in the DWT domain and the associated perceptual weighting function can be pre-calculated for each subband as shown in Fig. 4. Therefore, the complexity of CSF implementation in WTGM becomes the coefficient multiplication from the look-up table. This can be efficiently done in linear-time.

Regarding the complexity of NVF, $\eta(\gamma)$ and gamma function can be pre-calculated by the look-up table while the shape parameter is decided. $r(x, y)$ in Eq. (6) is determined by the local mean and the local variance which is related to the window size. The complexity of local mean and variance is $O(l^2)$, l $(= 2L + 1)$ is the window size. In this study, the window size is 3×3 for $L = 1$. Besides, the global variance is obtained for each wavelet subband and there are 12 subbands after 4 level wavelet decomposition. The total amount of calculation approximately equals to the image size (we can use static array to store the results). Thus, global variance takes $O(n^2)$ computation and the overall time complexity for NVF is no more than $O(n^2)$ $(O(n^2 \cdot l^2 + n^2) \approx O(n^2))$ since image width n is much larger than l.

Table 10 Summary of WTGM with WTQ schemes.

Operations	WTGM S_1	WTGM S_2	WTQ
Common Image Processing Attacks			
JPEG Compression (QF = 30%)	√	√	
SPIHT Compression (bitrate = 0.3)	√	√	
JPEG2000 Compression	√	√	N/A
Histogram Equalization		√	
Image Cropping (25% upper-left corner)		√	N/A
Brightness Enhancement (10%)	√	√	N/A
Contrast Enhancement (10%)	√	√	N/A
Median Filtering (4×4)	√		
Gaussian Filtering		√	√
Sharpening		√	
Rescale (50%, 90%, 150%)	√	√	N/A
Geometric Attacks			
Pixel Shifting (12 pixels)		√	
Rotation (2.5°)		√	
Security Measurement			
Multiple Watermarking (6 watermarks)	√	√	
Bitplane Removal (7 bitplanes)	√	√	

From our simulation, the whole loop of WTGM embedding and extraction under Intel Pentium 3.0 GHz, 1GRAM will need less than 2 seconds to complete for 512×512 testing images. In conclusion, the WTGM complexity is low and suitable for practical applications from the mathematical analysis and simulation results.

5.6 Summary

In general, the WTGM with relatively high-frequency components — WTGM(S_2) is superior to other methods, which can effectively resist common signal processing, geometric distortions as well as cryptanalysis. Also, it provides a better visual effect than other methods. The WTGM with relatively low-frequency components — WTGM(S_1) can be more effective in resisting JPEG and SPIHT compression as well as cryptanalysis, but ineffective in resisting geometric distortions. (as shown in Table 10).

In addition, WTGM does not use quantization to embed the watermarks and the cryptanalysis-like attack for WTQ is not useful to remove the watermark for WTGM. Therefore, we can clearly see that WTGM outperforms WTQ in almost all categories from the detailed comparsion. In general, the WTGM with medium-high frequency setting WTGM(S_2) is superior in resisting common signal process-

ing, geometric distortions as well as cryptanalysis with better visual perception than WTGM(S_1). Due to the difference of watermark embedding location for setting S_1 and S_2, the results are expected compared with other wavelet based approaches. However, the weakness for the WTGM is that the tree combination information must be kept secret which addresses extra storage space. The extended study should working on the design to efficiently reduce this extra cost

6. Conclusion

An efficient differential energy watermarking algorithm based on wavelet tree group modulation has been presented. In the proposed algorithm, the watermark is embedded in the relatively high-frequency components using the group strategy for each super tree such that energies of sub-super tree A and that of sub-super tree B are close. The employment of wavelet tree structure, sum-of-subsets and positive/negative modulation effectively improve the robustness of the watermark. The consideration to the CSF and NVF of the HVS provides a better visual quality of the watermarked image.

Compared with the WTQ scheme, the advantages of the proposed algorithm are as follows.

1) The proposed algorithm can tolerate more common signal processing and geometric attacks.
2) The length of the image key is large, which renders a better confusion/diffusion for security.
3) The human visual characteristics are considered in the wavelet tree based watermarking systems to provide a better visual quality.
4) The watermark can be public for users, and if any malicious user tries to destroy the watermark and sell those attacked copies, the user could be identified.

On the other hand, there are still some issues needed to be further studied as following:

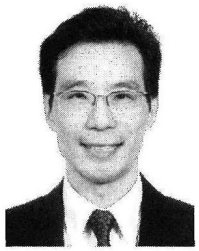
1) The group information for trees needs to be kept for watermark extraction, which needs more storage space.
2) The tolerance for geometric attacks is sill insufficient, feature-based or other RST (Rotation, Scaling and Translation) invariant mechanisms can be taken into account for better synchronization.

Acknowledgments

1. This work was supported by the National Science Council in Taiwan, Republic of China, under NSC95-2416-H009-027 and NSC96-2416-H009-015.
2. Partial technical background has been presented in the conference ICASSP 2007 [8].

References

[1] I.J. Cox, J. Kilian, F.T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Process., vol.6, no.12, pp.1673–1687, Dec. 1997.
[2] C.S. Lu and H.Y.M. Liao, "Multipurpose watermarking for image authentication and protection," IEEE Trans. Image Process., vol.10, no.10, pp.1579–1592, Oct. 2001.
[3] G.C. Langelaar and R.L. Lagendijk, "Optimal differential energy watermarking of DCT encoded images and video," IEEE Trans. Image Process., vol.10, no.1, pp.148–158, Jan. 2001.
[4] S.H. Wang and Y.P. Lin, "Wavelet tree quantization for copyright protection watermarking," IEEE Trans. Image Process., vol.13, no.2, pp.154–165, Feb. 2004.
[5] T.K. Das, S. Maitra, and J. Mitra, "Cryptanalysis of optimal differential energy watermarking (DEW) and a modified robust scheme," IEEE Trans. Signal Process., vol.53, no.2, pp.768–775, Feb. 2005.
[6] T.K. Das and S. Maitra, "Cryptanalysis of wavelet tree quantization watermarking scheme," IWDC 2004, pp.219–230, 2004.
[7] C.S. Lu, S.K. Huang, C.J. Sze, and H.Y. Liao, "Cocktail watermarking for digital image protection," IEEE Trans. Multimed., vol.2, no.4, pp.209–224, Dec. 2000.
[8] M.J. Tsai and C.H Shen, "Wavelet tree group modulation (WTGM) for digital image watermarking," IEEE International Conference on Acoustics, Speech and Signal Processing, vol.II, pp.173–176, 2007.
[9] J.L. Mannos and D.J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," IEEE Trans. Inf. Theory, vol.20, no.4, pp.525–536, July 1974.
[10] B.B. Huang and S.X. Tang, "A contrast-sensitive visible watermarking scheme," IEEE Multimedia, vol.13, no.2, pp.60–66, April-June 2006.
[11] L. Yong, L.Z. Cheng, and Z.H. Xu, "Translucent digital watermark based on wavelets and error-correct code," Chinese J. of Computers, vol.27, no.11, pp.1533–1539, Nov. 2004.
[12] A.P. Beegan, L.R. Iyer, and A.E. Bell, "Design and evaluation of perceptual masks for wavelet image compression," Proc. 10th IEEE Digital Signal Processing Workshop, pp.88–93, IEEE CS Press, 2002.
[13] M.J. Tsai and C.W. Lin, "Wavelet based multipurpose color image watermarking by using dual watermarks with human vision system models," IEICE Trans. Fundamentals, vol.E91-A, no.6, pp.1426–1437, June 2008.
[14] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," Proc. 3rd Int. Workshop Information Hiding, pp.211–236, Dresden, Germany, Sept. 1999.
[15] B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C, Second ed., John Wiley, New York, 1996.
[16] PGP, http://www.pgp.com/downloads/freeware/index.html
[17] P. L'Écuyer, "Uniform random number generation," Annals of Operations Research, vol.53, pp.77–120, 1994.
[18] JPEG 2000 compression, the International standard (IS 15444-1: JPEG 2000) published from ISO/IEC, [Online]: http://www.ece.uvic.ca/mdadams/hasper/
[19] StirMark, http://www.petitcolas.net/fabien/software/StirMarkBenchmark_4_0_129.zip
[20] I. Daubechies1 and W. Sweldens, "Factoring wavelet transforms into lifting steps," Journal of Fourier Analysis and Applications, vol.4, no.3, pp.247–269, May 1998.
[21] R. Sedgewick, Algorithms in C, Parts 1-5: Fundamentals, Data Structures, Sorting, Searching, and Graph Algorithms, 3rd ed., Addison Wesley, 2001.



Min-Jen Tsai received the B.S. degree in Electrical Engineering from National Taiwan University in 1987, the M.S. degree in Industrial Engineering and Operations Research from University of California at Berkeley in 1991, the Engineer and Ph.D. degrees in Electrical Engineering from University of California at Los Angeles in 1993 and 1996, respectively. From 1996 to 1997, he was a senior researcher at America Online Inc. In 1997, he joined the Institute of Information Management at the National Chiao Tung University in Taiwan and is currently an associate professor. His research interests include multimedia system and applications, digital forensic, digital watermarking and authentication, web services, enterprise computing for electronic commerce. Dr. Tsai is a member of IEEE, ACM, and Eta Kappa Nu.

Chang-Hsing Shen has received B.S. degree in information management from National Central University in 2000, the M.S. degrees in Institute of Information Management at the National Chiao Tung University in the year 2006. From year 2002 to 2003, he was in the development team of anti-virus service at Trend Micro. He later joined Inventec Besta in 2006 and focuses on digital entertainment and learning of mobile device.